Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.  
http://researchspace.auckland.ac.nz/feedback

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form.
Federated Searching and diverse subject areas in academic tertiary libraries: a study at database level

By Vanessa Joy Newton-Wade

Submitted to the School of Information Management, Victoria University of Wellington
in partial fulfilment of the requirements for the degree of Master of Library and Information Studies

June 2008
For Mum
Abstract

Searching online databases is a task tertiary students are commonly required to undertake when completing their academic work, and can often be a somewhat tedious process. Federated searching products appear to mitigate the tedium of this process by reducing the number of times a search needs to be repeated. This study aims to investigate the information needs of undergraduate history and medical students, what type of search results they prefer and their attitudes to the concept of a federated search engine.

Twelve students were interviewed, six from each subject. A semi-structured interview method was used for data collection. Results showed that the information needs of both groups of students are similar, as are the techniques they use to conduct and refine their searches. There was no correlation between the subject studied and the type of results preferred (specific or broad), but significant differences were found in the attitudes of the two groups of students to a federated search product. Implications of these findings on the suitability of a federated search in a tertiary academic library setting are discussed, along with the implications these findings have for information literacy. Suggestions are made for further research.

Keywords: Information needs, Federated Searching, Tertiary Libraries, Semi-structured interviews.
Acknowledgements

This project represents the end of a two and a half year journey. I would like to extend my thanks to all of those who have helped me along the way.

Special thanks go to my supervisor, Brenda Chawner, for her expertise, patience and making time in her busy schedule to meet with me when she was in town. It can't be easy supervising from a distance!

My appreciation to the History Department and the Faculty of Medical and Health Sciences at The University of Auckland for permitting this research to be undertaken in their departments. Thanks also to the staff of the Philson Medical library for their assistance and willingness to help with recruiting participants.

The University of Auckland Library has provided assistance on many levels (not least for keeping a roof over my head and food on the table). Thank you to my colleagues for assistance and stimulating conversation, particularly Leonie Hayes and YinYin Latt, and for the time off work when needed to complete this project.

Thanks also to the Tofu Elephant Club for the support, lunches and most of all the laughs. Kudos to my friends (thanks for proofreading Liza), family and my partner Anton for putting up with me while I was glued to my laptop.

Most of all, thank you to those students who took time out of your schedules at a very busy time of semester to participate in this project. Thank you - it would not have been possible without you.
IMPORTANT DISCLAIMER
with respect to a MLIS Research Project (INFO 580)

‘Federated Searching and diverse subject areas in academic tertiary libraries: a study at database level.'
(hereafter referred to as 'The MLIS Research Project')

being undertaken by
Vanessa Joy Newton-Wade

in partial fulfilment of the requirements of the degree of Master of Library and Information Studies, School of and Information Management, Victoria University of Wellington.

Topic Commencement: **December 2007**

1. Victoria University of Wellington and its Council, its members, staff, employees, students and agents undertake no duty of care in contract, tort, or otherwise, to users (whether direct or indirect) of the MLIS Research Project and make no warranties or representations of any kind whatsoever in relation to any of its contents.
2. The MLIS Research Project is only made available on the basis that all users of it, whether direct or indirect, must take appropriate legal or other expert advice in relation to their own circumstances and must rely solely on their own judgement and such legal or other expert advice.
3. Under no circumstances will Victoria University of Wellington and its Council, its members, staff, employees, students or agents be liable in any way whatsoever, whether in contract, tort (including negligence), for breach of any statutory or regulatory duty (to the fullest extent permissible by law), or otherwise, to any user (whether direct or indirect) of the MLIS Research Project for any loss or damage whatsoever arising directly or indirectly as a result of the use in any way of the MLIS Research Project.
4. Each exclusion in the clauses of this disclaimer and each protection given by it is to be construed as a separate exclusion applying and surviving even if for any reason any of the exclusions or protections are held inapplicable in any circumstance.
# Table of contents

Abstract ..................................................................................................................................... i
Acknowledgements .................................................................................................................. ii
Table of contents .................................................................................................................... iv
1 Introduction ...................................................................................................................... 1
2 Literature Review ............................................................................................................. 2
  2.1 Federated Searching ........................................................................................................ 3
  2.2 Background and development of federated searching .................................................. 3
  2.3 Studies involving end-users ......................................................................................... 4
  2.4 Methodologies ............................................................................................................... 4
  2.5 Opinion-based articles: Why to choose federated searching ........................................ 5
  2.6 Evaluation literature .................................................................................................... 6
  2.7 ‘Federated Searching and Google’ ................................................................................ 6
  2.8 Comparison of federated searching products ................................................................ 7
  2.9 Literature written by vendors ..................................................................................... 7
  2.10 Architecture ............................................................................................................... 8
  2.11 Conclusion: where are the gaps? ............................................................................... 9
3 Objectives ........................................................................................................................ 9
4 Research Questions ........................................................................................................... 9
5 Researcher Bias ............................................................................................................... 10
6 Research Paradigm ......................................................................................................... 10
7 Population ....................................................................................................................... 11
8 Methodology ................................................................................................................... 12
9 Results ............................................................................................................................ 15
10 Discussion ..................................................................................................................... 24
11 Directions for further research ...................................................................................... 30
12 Conclusion .................................................................................................................... 31
13 Bibliography ................................................................................................................. 33
14 Appendices ................................................................................................................... 36
  14.1 Participant Information Sheet .................................................................................... 37
  14.2 Participant Consent form .......................................................................................... 38
  14.3 Interview questions: ................................................................................................. 39
  14.4 Sample transcript ...................................................................................................... 40
14.5  Sample notice from Cecil ................................................................. 41
14.6  Sample poster .................................................................................. 42
1 Introduction

Libraries in New Zealand are just beginning to evaluate and implement federated searching technology, and there has been no formal research done in New Zealand to date about the effectiveness of federated searching.

The aim of this study is to extrapolate user attitudes to the concept of federated searching from interviews with tertiary History and Medical students about their searching behaviour, and what their searching needs are. Federated searching claims to provide an effective ‘one stop shop’ for users to search across a variety of databases at once. However, librarians are ‘concerned that searching might only work at the lowest common denominator’ (Curtis & Dorner, 2005), and this concern has prompted investigation at a local level. There are a limited number of live federated searching sites across the library sector, but implementation of federated searching has not yet reached a critical mass in this country. This will change in 2008 – in the tertiary sector, Auckland University of Technology (AUT) has already implemented 360 Search (Solutions, 2008), known at AUT as Multisearch, and others in the Library Consortium of New Zealand will also implement 360 Search this year (LCONZ, 2007). Other tertiary institutions outside of the consortia are likely to follow suit with federated searching technology in one form or another (pers. comm., 2007).

Across the public library sector, federated searching is the exception rather than the rule. Wellington City libraries offers a federated searching solution through WebFeat, but an informal investigation of the electronic resources page for the other major New Zealand public libraries (Auckland, Christchurch & Dunedin) showed that no other public libraries are using federated searching as yet.

The genesis for this research resulted from the limitations of studying federated searching technology directly. Given the limited number of live federated searching sites available in New Zealand, and the limited amount of time this technology has been available to New
Zealand users, it was decided not to study the technology directly. The research problem in a climate where libraries are beginning to implement a new technology is a logical progression from researching federated searching directly. As it was not possible to study the federated search technology directly, it was decided to take a ‘step back’ and investigate the needs and attitudes of potential users. As researching federated searching technology directly in New Zealand is not yet practical due to a limited number of live installations, an opportunity exists for research on this topic in the future.

As researching the entire library sector is outside the limited scope of this study, the focus has been narrowed: first to an academic tertiary institution’s library and then further, to students of two contrasting academic disciplines: history and medicine. This study will evaluate the attitudes and experiences of tertiary history and medical students to database searching and extrapolate the potential impact of federated searching on students of these subjects from the data collected.

Several basic questions stand out, and form the basis for this research: Is a federated search an appropriate tool for specialist academic research? Do students of diverse academic areas search in the same or different ways? What are their attitudes to the concept of federated searching?

2 Literature Review

This literature review has focused on federated searching, although in the context of the research being undertaken a section on user needs may not have been unexpected. However, the relevant material on this topic has already been covered in detail in other sections in the proposal (theoretical framework and methodology).
2.1 Federated Searching
Federated searching has been available on the World Wide Web for some years, although libraries have been comparatively slow to pick up on this technology. Curtis & Dorner (2005, p.35) define federated search as a type of search that searches across multiple databases, returning organised results to the user. The literature available is sharply delineated into two types: IT-based and Library and Information Management-directed. As this is a new and growing field in the Library and Information Management sector, the literature available is somewhat limited. Every effort has been made to procure as much Library and Information Management literature on the topic as possible, as well as a selection of the more well established Information-Technology directed literature.

The literature available charts the development of federated searching from world-wide-web engines and enterprise level to the development of a library-specific product.

Items selected for this review were restricted to articles or chapters primarily about federated searching. There is a body of literature focussed on serials that mentions federated searching, but not as the primary subject under discussion. The literature located for this review covers those topics in more detail. The literature discussed has all been published within the last three years. When searching, date limits were placed to filter out publications before the year 2000. Very few articles were found before 2003, and none that addressed federated searching in a library context in detail, other than to outline it as an ideal goal.

2.2 Background and development of federated searching
Much of the available literature is introducing the concept of federated searching to librarians and discussing the development of the product. As a new technology that until recently has been seen as difficult both to use and to implement, recent advances have led to an explosion in writing on this topic.

Fryer (2004) and Tennant (2005) provide a good explanation of what federated searching is, how it functions and what the major products available are. De-duplication of results is a
topic that appears across the board – this is clearly a new development and a solution to a problem that is identified in the earlier literature. Joy and Sennyey (2006) provide a place for federated searching as one of many services provided by a library, and look at the costs involved in providing what was seen as a ‘value-added’, extra service at this time.

2.3 Studies involving end-users
Formal academic research on federated searching is just beginning at this point in the development of the technology. In order for effective research to be carried out, a critical mass of ‘live’ sites running the technology for a reasonable length of time needs to be available. Reeb et al. (2006) interviewed undergraduate students about their experiences using federated searching and the results shed light not only on the federated search tools being evaluated, but also the searching behaviour of undergraduate students. Hsieh-Yee, Tang and Zhang (2006) took a different approach, studying user perceptions of federated searching via a simulation of MetaLib. The full results of this study were not available at this time, however they promise to be interesting. Preliminary results showed a marked difference in the way library staff viewed federated searching being used and the way it was used by students. Interestingly, a significant number of librarians (57%) found federated searching difficult to use compared with novice searchers (8%). More of this type of literature will become available as more sites implement the technology.

2.4 Methodologies
Hsieh-Yee, Tang and Zhang (2006) and Reeb (2006) have all done studies involving end-users and their perceptions and usage of federated searching. Hsieh-Yee (2006) took cohorts from three groups and compared survey data using a simulated search. Reeb (2006) conducted end-user interactive interviews using real-time observed searching of a library catalogue and databases. Reeb (2006) acknowledged difficulties with this method, including the need to point users to federated technology if they did not make their way there spontaneously. Search techniques can be difficult to study, and are made more difficult by
this particular methodology choice, in particular the need to point users to federated technology. A simulated search, such as that used by Hsieh-Yee, provides more reliable and unbiased results.

Chen (2006) tested MetaLib and WebFeat at several live locations, running test searches and comparing the results. This type of methodology allows a direct comparison between two types of federated searching, allowing libraries to make informed choices backed up by formal research.

2.5 Opinion-based articles: Why to choose federated searching
While these are not academic studies, in a new and emerging area such as federated searching, opinion-based articles provide valuable clues about where gaps are in the research. If an issue is being discussed by librarians, chances are that it is interesting or important to the sector. If there is no research on these topics as yet, then these opinion-based articles can provide a good indication of where the ‘gaps’ lie.

There is plenty of discussion about federated searching among librarians. The initial layer of discussion on a new technology relates to the perceived efficiency and usefulness of it. In short, why use (in this case) federated searching over the more conventional OPAC searches available? Fahey (2007), Curtis and Dorner (2005) and Tennant (2003) provide three different approaches to discussion about federated searching. Tennant (2003) provides a very good overview of how federated searching works in the library sector.

Fahey (2007) provides insight into the more philosophical debates librarians are having over federated searching – that it ‘dumbs down’ the searching process and that students miss out on important information. Curtis (2005) provides a concise description of federated searching and also makes the point that it has been designed for a hybrid library: to search an OPAC catalogue and online databases simultaneously. This is an emerging aspect of federated searching: it is only recently that hybrid searching has been a viable option.
2.6 Evaluation literature
A large proportion of the literature, including some that have already been discussed as ‘opinion’ works can be categorised as ‘evaluation’ literature. That is, it exists to facilitate decision-making by librarians about whether a new type of product is likely to suit the library involved. Much of this literature is not formally research-based, but describes the implementation and experience of a library or institution with a new product or service.
Newton and Silberger (2007) discuss the implementation of Central Search from Serials Solutions.

Highsmith and Pontford (2006) describe the implementation of MetaLib at Texas A & M University. They outline the process from selection of a product through to implementation and training of staff. Boyd et al. (2006) take a range of short descriptions of American libraries’ experiences evaluating and implementing various types of federated searching.

Cervone (2005) provides a detailed analysis of the usability of several federated search products, discussing issues and potential solutions, backed up with external research.

2.7 ‘Federated Searching and Google’
Google is an obvious point of comparison for federated searching outside library vendor’s products, for a number of reasons. Google and Google Scholar are perceived by librarians to be the ‘ideal’ search function of choice for many patrons. As a result, libraries are striving to provide a ‘single search box’ for patrons – federated searching is the closest we have come thus far. Okerson (2006) provides an alternative view of Google. Although this article does not seem to have much to do with federated searching at first glance, focusing on digitization projects, it does provide a good insight into the attitudes of librarians about Google. Okerson argues obliquely that perhaps instead of developing software that replicates the ‘Google’ function, perhaps Google should be brought onside and a cooperative searching solution developed. Newton and Silberger (2007) provide a step along this path: Google is included in a federated search of the institution’s databases. This has not yet been replicated at other
institutions, but has apparently proved popular with students. No quantitative data was provided to back up this statement: the data was gathered through informal feedback from students. Nonetheless, combining Google with a federated search of library databases is logical from a user’s perspective and one that is worth investigating further. However, it does raise questions about how results are ranked from such a search.

2.8 Comparison of federated searching products
Boss (2005) and Chen (2006) make a comparison of two or more commercially available federated searching products. This type of research is particularly valuable, as they fill a gap in the literature. Chen (2006) compares WebFeat, MetaLib and Google. This involves looking at the products from the perspective of both users (although no direct user interviews are made) and system administrators.

Pesch (2006) evaluates the updated version of WebFeat that has just been released. The addition of hosted searching is an innovation that impressed Pesch. This allows libraries without access to onsite technical expertise the ability to offer federated search technology, enabling access to this technology for smaller libraries that may be operating under tight budget constraints, or with limited resources.

2.9 Literature written by vendors
There is a small body of literature published in academic journals that has been written by authors with affiliations to vendors or publishers, usually employees. Marshall, Herman and Rajan (2006) are all employed by Swets Information Services. Their research looks at federated searching in the context of the databases provided by Swets. This article has a sound academic base, drawing on many of the articles already mentioned in this literature review. A detailed discussion of the features Swets offers that work well with federated searching follows. The depth of the referencing and discussion propels this article beyond a glorified advertisement for a vendor’s product. Noerr (2006), on the other hand, describes Muse Metasearch without the detailed referencing that the Marshall article offered. As a
result, the Muse Metasearch article has more of an overt bias. Noerr does include a useful
description of what makes a good metasearch ‘publishing’ environment.

No vendor websites have been included in this literature review, as the material found on
these sites is not detailed enough to warrant it.

2.10 Architecture
There is a large body of literature available about the technical details of federated searching
architecture and the development of federated searching. Solomon (2004) provides a slightly
earlier look at federated searching. The article bridges the gap between the Library and
Information Management and Information Technology literature in that it covers many of the
issues discussed in other articles in this review (particularly issues discussed in opinion-
based articles), but in a more technical manner. Si (2003) provides insight into the
background and technical difficulties of merging search engine results, presenting a new
solution to these difficulties. This article is very technical, but provides insight into the inner
workings of search algorithms.

Trnkoczy, Turk and Stankovski (2006) provide details of a grid-based architecture developed
for personalised federation of digital libraries. This architecture is based on the OAI-PMH
protocols and harvests the metadata, creating an index, rather than collecting the digital
items themselves. This allows a user to create a personalised federated library tailored to
their individual research needs. This article indicates a possible future direction for federated
searching.

Avrahami (2005) discusses a federated search system developed for the US government.
There are many lessons libraries can learn from the federated searching technology
developed in this type of situation: this search function searches 20 government websites
simultaneously. The level of detail that this article goes to in order to describe the search
functionality involves some descriptions of algorithms, ranking criteria and result merging.
2.11 Conclusion: where are the gaps?
There are a number of large gaps in the literature. There is no writing on New Zealand experiences of federated searching, for the simple reason that there are very few, if any, implementations in this country. This may change over the coming months, and if it does there is certainly scope for research in this area. On a broader plane, there has been little academic research done on federated searching as yet. This is not surprising, given that this is an emerging technology. There have been a few end-user studies, but until the number of sites utilising federated searching increases the number of studies is unlikely to increase much. There have been a number of comparisons of vendor products. Surprisingly, although there are not yet many implementations of the new product advertised by Ex Libris, ‘Primo’, there has been very little literature written about it. Once this product has more live sites, there is scope for research assessing the product and its impact on the libraries which have installed it.

3 Objectives
The objectives of this study are to identify situations where federated searching may be appropriate for academic tertiary libraries, and to discover whether there are innate differences in the preferred searching methods used by students of contrasting academic disciplines.

4 Research Questions
What are the information needs of history and medical students when searching online databases?

How can the desired type of search result (e.g. precise ‘hit’ vs. broader result) affect the effectiveness of a search?

What are the attitudes of history and medical students to a product that will search more than one database simultaneously, given the limitations of this type of searching? Is there a correlation between the subject studied and the type of search preferred?
5 Researcher Bias

The subjects for this research (History and Medicine) were deliberately chosen to minimise researcher bias, while still allowing the researcher to have some knowledge of each field. The researcher’s primary subjects at tertiary undergraduate level were Music and English, although one History paper was taken at first-year undergraduate level. The researcher has a year’s experience working in an academic medical library. As a result, the researcher has a good knowledge of how students in both subjects go about using their respective library resources.

6 Research Paradigm

A post-positivist paradigm has been adopted, and a semi-structured interview process was selected to allow a consistent approach, but also allow greater flexibility than a fully structured interview. Combined with contextual enquiry, this should provide a good level of detail in the results. It should be noted that post-positivism and interpretivism share some characteristics, and if this study were placed on a continuum it would sit closer to interpretivism that positivism. The reason a post-positivist approach has been chosen instead of an interpretivist approach is because this study seeks to explain the ‘how’ (understanding) and the ‘why’ (explanation) of searching behaviour, essentially attempting to bridge the gaps between positivism (explanation) and interpretivism (understanding) (Grix, 2004).
7 Population

The population for this study is undergraduate students at the University of Auckland, New Zealand. The sample used was six students from two disciplines, History and Medicine.

Limitations on the sample included:
First year undergraduate students were not interviewed as they are generally not expected to use online databases to the same extent as second year & above. As this study was carried out in the first semester of the academic year, it is unlikely that first year students would have been familiar enough with the library resources to provide detailed responses to the interview questions.

Undergraduate students only were interviewed – there is scope for further research using a postgraduate cohort, however, the small sample size meant skewing was mitigated by restricting the study to undergraduate students.

Students were self selected – volunteers were called for via announcements in year two & three History classes, word of mouth, two postings on the online learning software that all medical students access on a regular basis ('Cecil'), and verbal announcements over the PA system in the Philson Medical Library at the University of Auckland. Research participants were also asked to spread the word about the research to their friends.

The different canvassing methods used for the two subjects places some limitations on the study – medical students who used the Library were more likely to hear about the study, for example. Recruiting methods that had particular success were announcements made in History classes, particularly one class where a sign-up sheet was passed around during class, and the announcements put on Cecil for the medical students. Students who participated in the research were offered a voucher for a free coffee at a local cafe by way of thanks.
Using volunteers also places some limitations on this study – a random sample was
unachievable as a result of using volunteers. Interviewees were less likely to participate
unless they have an interest in doing so. The small sample size does not allow
generalisation of results, and using volunteers means a level of bias is inherent in the
results. This does not affect the validity of the study, but does mean that the results may
differ from the same study conducted using a random sample. Depth of data and detail is
important to achieving the objectives of this study, and the sample size combined with the
methodology described below will achieve this.

8 Methodology

This study was undertaken using qualitative methodology, which allows much more freedom
to follow lines of enquiry that are more appropriate to this situation than a qualitative study
would allow.

A semi-structured interviewing technique was chosen as the data collection methodology for
this study. Interviewing allows a narrow, but very detailed view of the chosen subject area to
be built up, in the sense that a large sample size is not always necessary. Interviews are
used when ‘the nature of the data is too complicated to be asked and answered easily’
Pickard, (2007). An entirely structured interview would not allow the same degree of
flexibility in terms of the responses received, and in the opinion of the researcher, delivers
little that a well-structured survey or questionnaire could not. Grix (2004) acknowledges this:
‘the drawback is that this technique [structured interviewing] is inflexible and not designed to
cope with the unexpected’. The data delivered by such an interview would likely not be an
accurate indication of what the research participant’s needs are, unless the interview
questions are extremely well-designed and scoped. By allowing the interviewer to engage
with the research participants and follow up unexpected lines of enquiry the data returned
was of greater depth than other data collection methods, such as a questionnaire or survey.
Berg (2004) states that ‘interviewers are permitted (in fact, expected) to probe far beyond the
answers to their prepared standardized questions’. Allowing flexibility in the interview situation increases the level of detail in the data gathered, and allows a much clearer picture to be built up of what the research participants require from federated searching technology.

Nonetheless, as with any data collection method, there are drawbacks. Usability testing, combined with semi-structured or unstructured interviewing of research participants would probably provide the most detailed and accurate record of how research participants search online databases, and what their ‘wants’ and ‘needs’ are from future searching products. However, semi-structured interviews combined with contextual enquiry are most appropriate methodologies due to the following limitations on this study: timeframe, facilities available, and cohort.

Human Ethics Approval from the School of Information Management at Victoria University of Wellington was sought and granted prior to seeking participants. Permission from the Heads of Departments at the Faculty of Medical and Health Sciences and the History Department in the Faculty of Arts at the University of Auckland to interview their students was also granted prior to seeking participants. Participants were contacted via the methods discussed in the Population section above, and were either sent a participant information sheet ahead of time via email, or asked to read one before commencing the interview depending on the method of initial contact. Participants were required to sign a consent form prior to beginning the interview (see the Appendix for copies of the consent form and participant information sheets). Included on the consent form was an option to request a copy of their interview transcript, and an option to receive a summary of the research results. Participants were given a period of two weeks from the interview date to pull out of the research if they wished. No participants chose to pull out.

Each of the 12 participants was asked seven set questions, after being given a scenario to consider before they answered the questions (see below; a copy is also in the Appendix).
Depending on the responses received, different lines of enquiry were followed up in order to generate a coherent and homogenous data set.

Starting point: Think about, or imagine, beginning research on a topic you know little about, using a library database you are familiar with. You are not searching for known references (e.g. an article or author you have a reference for). An example might be beginning to look for information on an essay topic.

1. How do you go about your initial search?
2. What do you consider a successful search?
3. What would you do differently if your search is unsuccessful?
4. What is your tolerance limit before ‘giving up’ on a database and trying something else? What are your typical reasons for ‘giving up’ on a database?
5. Do you have a preference for results that are specific or broad in terms of the number of results returned and their relevance, and why? How specific do you need the information to be at this point in your research?
6. What is your reaction to a search engine that can search all of the library databases from all subjects at once?
7. Is there anything else you would like to tell me about how you search for information?

Participants were interviewed in front of a computer or laptop with the University of Auckland’s library page open on the screen, and participants were told that if they wished to search during the interview they were welcome to. The purpose of this was to allow participants context to inform their answers, if they wished. Not all of the participants chose to search databases as the interview was taking place. The reason for this is that interviews took place between weeks nine and eleven of a twelve-week semester, at a busy time leading up to exams. Many assignments are due at about this time of semester and all participants could think of a recent assignment or essay topic that fit the scenario with no problems. Some could recall recent searches in detail, whether they were successful searches or not, and used this experience to inform their answers.

In terms of interview technique and rapport, the exact questions participants were asked were refined as the interviews progressed. All of the set questions were the same, however, the follow-up and probing questions that participants were asked developed over the course of the interviews. The lines of enquiry were also further developed as the interviews progressed. Responses given by early interviewees informed some of the lines of enquiry taken in later interviews, for example further questions on the concept of faceted searching.
An effort was made on the part of the interviewer to keep the tone of the interviews friendly and develop a rapport with the interview participants. It was felt that by keeping the tone of the interview more casual, that participants would feel more comfortable with the interview process, and that the data gathered would be a more accurate reflection of how participants thought and how they search.

Interviews took between 6 and 13 minutes, depending on the depth of answers received and the lines of enquiry followed. Interviews were recorded digitally in .wav format, and transcribed (see the Appendix for an example of the transcription). General, overarching themes were then drawn out from the data via close reading of all transcripts, and open coding was undertaken in order to analyse the data more closely. The coding was undertaken firstly on a question by question basis, with responses examined to draw out similarities and differences between the two groups interviewed. Axial coding was then undertaken to understand the ‘when, where, what, and how’ (Strauss and Corbin, 1998, p.18), and to provide more depth and detail of results. Broad themes were also drawn from the data, these are referred to in the discussion section below.

9 Results

As discussed in the methodology above, semi-structured interviews were used for the data collection. The structured portion of the interviews consisted of six set questions, and most participants also answered several similar follow-up questions. This section will set out the results from each of the ‘structured’ questions, which all participants answered, followed by the ‘unstructured probing’ questions relating to the set question. The tables following each interview question contain the raw coded data. It is important to note that a single response may have been coded more than once, depending on the categories the response fell in to. Where an entry in a table has a number in brackets, this indicates the number of responses on this theme. Where the discussion of results refers to proportions (e.g. half) of the cohort, a single response only was coded from each participant. Some of the follow up or probing
questions have not been tabulated as only one or two responses in total were received.

These responses have been quoted where relevant. Participants are referred to by their subject, followed by a number e.g. Medical participant #1. Following this section, the overall results and broad themes drawn from examining the data as a whole will be covered in the Discussion section that follows.

The initial scenario given to participants was to

“Think about, or imagine, beginning research on a topic you know little about, using a library database you are familiar with. You are not searching for known references (e.g. an article or author you have a reference for). An example might be beginning to look for information on an essay topic.”

The intention was to limit the study slightly – the way searches are conducted differs depending on the aim of a search (e.g. ‘hunting down’ the fulltext of a known citation, as opposed to a more generalist information-seeking search). By asking participants to focus on the beginning stages of the information search, it was hoped to focus on a less rigidly structured search, resulting in a more detailed data set.

1. How do you go about your initial search?

<table>
<thead>
<tr>
<th>History Cohort</th>
<th>Medical Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>JSTOR (x3)</td>
<td>MEDLINE (x3), PubMed (x2, both 5th yr), Scopus x1, PsychInfo x1</td>
</tr>
<tr>
<td>Keywords (x4)</td>
<td>Keywords (x3)</td>
</tr>
<tr>
<td>Google (x2, x1 specifically mentioned using Google 1st)</td>
<td>Google (x4, x3 specifically mentioned they go to Google 1st)</td>
</tr>
<tr>
<td>Prior experience (x4)</td>
<td>Prior experience (x4), trial &amp; error (x1)</td>
</tr>
<tr>
<td>LEARN (University of Auckland Library gateway), (x2)</td>
<td>Google Scholar (x2)</td>
</tr>
<tr>
<td>Subject resources (via LEARN x1)</td>
<td>Depends on topic (x1)</td>
</tr>
<tr>
<td>Books (x1)</td>
<td></td>
</tr>
</tbody>
</table>

Responses to this question varied, depending on how participants interpreted the question. Some began by outlining their searching method by describing their route to a particular database, others at the print book level, and some described their searching methods within a particular database. All respondents discussed their searching methods within a database.

Just over half of the respondents from both subjects named a database they preferred to use in response to this question. JSTOR was the preferred database for half of the History
students interviewed, and Medline had the same level of preference for the Medical Students. Interestingly, twice as many medical students mentioned Google in response to this question than history students. Of the four medical students who mentioned Google, three explicitly stated that Google was their preferred starting point for their initial search. Google was used, in the words of a medical participant, ‘to make sure I know the right meanings of the words of the topic, to get an overall general idea’. All participants who mentioned using Google for their initial search used Google in this way.

All used keyword searches, with the keywords usually coming from the assignment topic they had selected. Some participants mentioned that they might also start their information search from a recommended reading list or bibliography given to them with their assignments.

All students stated at some point in their interview that their search techniques were the result of prior experience, or ‘trial and error’. Many students had developed a ‘favourite’ search method, a process they followed for each topic.

2. What do you consider a successful search?

<table>
<thead>
<tr>
<th>History Cohort</th>
<th>Medical Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant</td>
<td>Something useful, a start (or similar - x4)</td>
</tr>
<tr>
<td>1-2 useful articles</td>
<td>Results that answer my question (Relevant - x2)</td>
</tr>
<tr>
<td>4-5 quality articles</td>
<td>‘I’m not expecting much’</td>
</tr>
<tr>
<td>2-3 articles (if a narrow topic)</td>
<td>Not too many results (50ish)</td>
</tr>
<tr>
<td>5-6 articles (if a broad topic)</td>
<td>‘[something] that gets me some marks’</td>
</tr>
<tr>
<td>Reviews of books (indicative of quality of book)</td>
<td>Find what I’m looking for</td>
</tr>
<tr>
<td>Something to start with</td>
<td>Very specific</td>
</tr>
</tbody>
</table>

There were no discernable differences between the answers given by students of different subjects to this question. Naturally enough, all participants wanted relevant results that answered their particular question:

‘a successful search is... specific, and gives the search results that are able to answer my question’ (Medical participant #3).
‘... [results] that deal quite directly with the subject topic, that offers quite a lot specifically about the essay’ (History Participant #4).

What was noticeable was that four of the six participants in the history cohort mentioned or indicated a desired number of articles in response to this question. This was borne out by responses to other questions – the history cohort seemed to look for a certain number of references for their assignments more often than the medical cohort, none of whom mentioned a specific guideline for the number of references they were seeking. Several participants mentioned in answering other questions that time was usually an important factor, correlating the findings of Foster and Gibbons (2008, p. 59): ‘we learned how busy our students lives are...’, and this may well be reflected in the attitudes of participants to the database searching process.

The expectations of students from their initial search was varied, with most statements in response to this question outlining ‘a start’ (5 participants), or ‘something relevant’ (6 participants). As mentioned above, most of the history participants mentioned a specific number of articles, whereas the medical participants were more abstract in their responses.

3. What would you do differently if your search is unsuccessful?

<table>
<thead>
<tr>
<th>History Cohort</th>
<th>Medical Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change keywords before doing anything else (x6)</td>
<td>Change keywords before doing anything else (all 6)</td>
</tr>
<tr>
<td>Advanced search, add limits (x2)</td>
<td>Advanced search, combine terms differently in Medline, limits, etc. (x4)</td>
</tr>
<tr>
<td>Redefine search (x2)</td>
<td>Google (x2)</td>
</tr>
<tr>
<td>Look at bibliography for any relevant articles already found (x2)</td>
<td>Change my topic! (x2)</td>
</tr>
<tr>
<td>Try another database</td>
<td>Synonyms</td>
</tr>
<tr>
<td>Books/encyclopaedia (x2)</td>
<td>Ask a librarian</td>
</tr>
<tr>
<td>Widen/narrow search depending on results</td>
<td>Take a day off</td>
</tr>
<tr>
<td></td>
<td>Widen scope</td>
</tr>
</tbody>
</table>

If the initial search did not bring back the desired results, all participants adjusted their keywords and tried another search using the same database. Half of the medical cohort referred to usage of the ‘combine’ function in Medline, which is a powerful tool for narrowing down a search very quickly. These participants also adjusted the way in which
they combined their search terms, often before changing keywords. Many of the responses to this question related to various techniques participants used to check and/or change their keywords in order to improve the relevance of the results returned. These ranged from using Google, to encyclopaedia (both print and online), to using synonyms or different spellings of search terms. Six participants mentioned using some form of advanced search or limits function, and one mentioned widening or narrowing her search depending on the type of results she received. Interestingly, only one participant in this cohort mentioned asking for help from a librarian. Some participants gave very frank and honest answers, admitting that they might ‘take a day off’ before trying another search on the topic in question. Two medical participants admitted that if their searches did not provide the information that they needed that they might change their topic. One medical participant went on to explain further in a response to the next question:

‘...usually when we do research, it’s for some kind of literature review, and if it’s a free-range topic you can choose one. I usually give up [if I can’t find what I want] and then choose another topic. It’s an easy way out!’ – Medical Participant #6.

4. What is your tolerance limit before ‘giving up’ on a database and trying something else? What are your typical reasons for ‘giving up’ on a database?

<table>
<thead>
<tr>
<th>History Cohort</th>
<th>Medical Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 searches, 10 searches = frustrated!</td>
<td>Pretty quick – 3 diff searches</td>
</tr>
<tr>
<td>1 hour</td>
<td>Pretty high, usually get results wanted</td>
</tr>
<tr>
<td>So many databases avail, 15 min, then try something else</td>
<td>20min – half an hour</td>
</tr>
<tr>
<td>3x, no results, move on</td>
<td>Pretty patient – 3 days</td>
</tr>
<tr>
<td>Pretty short, impatient</td>
<td>10 searches</td>
</tr>
<tr>
<td>Dep on results, if relevant more patient</td>
<td>Only use Medline &amp; Pubmed so don’t give up</td>
</tr>
<tr>
<td>Time constraints</td>
<td>‘depends on the day’</td>
</tr>
<tr>
<td>Dread</td>
<td></td>
</tr>
<tr>
<td>Frustration</td>
<td></td>
</tr>
</tbody>
</table>

Subquestion – is the speed of a database a contributing factor to moving on?

<table>
<thead>
<tr>
<th>History Cohort</th>
<th>Medical Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>If very slow</td>
<td>If very slow</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Maybe (x2)</td>
<td>Doesn’t influence decision to move on</td>
</tr>
<tr>
<td>Not an issue (x2)</td>
<td>Not an issue (x3)</td>
</tr>
<tr>
<td>Definitely</td>
<td>Definitely</td>
</tr>
<tr>
<td>Yes, if doing a quick search</td>
<td></td>
</tr>
</tbody>
</table>

This question elicited the most emotional responses from participants. It was noticeable that on the whole, the medical cohort displayed more patience before moving on from a database than the history cohort. This is likely to be to do with the number and type of databases available to each subject – one 5th year medical participant only used Medline and PubMed, and tended not to give up on these sources as he knew the information he needed was there and that it is a matter of finding it. The other 5th year medical participant interviewed concurred with this response by stating that she usually got the results she wanted.

All participants gave either a rough indication of the number of searches that they might perform unsuccessfully, or the amount of time they were prepared to spend persevering unsuccessfully with a database before moving on and trying a different database or other type of resource (for example, Google or print books). Again, there were no correlations between indicated levels of perseverance between participants of different subjects.

There are a larger number of databases available to history students at the University of Auckland than to the medical students. The nature of history as a subject means that material used for research is not as clearly delineated, and more interdisciplinary than that required for scientific subjects, on the whole. The history cohort certainly displayed a higher level of emotional responses than the medical cohort when asked about their tolerance limits, as shown in the coding tables above. One history participant stated that she

“...dreads, I actually don’t look forward to having to go through the databases and article searching process, because that just feels overwhelming sometimes...”. History Participant #4
Several participants commented that database content was an influencing factor in moving on from a database. If after several searches the required information (or at least, information in the same or similar area) was not found, then an assumption might be made that perhaps the database being used was not quite right for the topic. A medical participant commented that

‘... sometimes you might guess that the database doesn’t have exactly what you are looking for... I’ll say maybe this will be more of a PsychInfo question, maybe this database isn’t covering all of the information I need it to cover, which is why I’m not coming up with the results I’m looking for’  Medical Participant #1

Probing questions about the database speeds were put to all participants, with the majority of respondents stating that the speed of a database was either not an issue, or not a problem unless very slow when deciding to move on from a database. In addition, some participants commented that if a database appeared slow, they were more apt to lay the blame with the overall internet speed, not with the database individually.

5. Do you have a preference for results that are specific or broad in terms of the number of results returned and their relevance, and why? How specific do you need the information to be at this point in your research?

<table>
<thead>
<tr>
<th>History Cohort</th>
<th>Medical Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad (x2), to begin with (x1: 3 total)</td>
<td>Broad (x2)</td>
</tr>
<tr>
<td>Specific (x3)</td>
<td>Specific (x4), depends on topic (x1: x4 total)</td>
</tr>
<tr>
<td>‘I’m not often in a position where I have lots of time to spend doing broad reading’</td>
<td></td>
</tr>
<tr>
<td>‘I’m sure I would have appreciated broader results in 1st year’</td>
<td></td>
</tr>
</tbody>
</table>

This was a reasonably subjective question, and the answer may be dependent on individual learning styles, rather than which subject a student studies. Although more medical participants than history participants stated that they preferred a specific search over a broad search, the margin was not high. Nonetheless, given the small sample size, that there is any difference at all is worth further investigation.
The amount of time available to participants for their research coloured some of the responses, with some participants commenting that if more time were available then they may prefer a broader search: ‘I’m not often in a position where I have lots of time to spend doing broad reading’. Some also commented that the level of specificity they required depended on the topic they were researching at the time.

6. What is your reaction to a search engine that can search all of the library databases from all subjects at once?

<table>
<thead>
<tr>
<th>History Cohort</th>
<th>Medical Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great/fantastic (x3)</td>
<td>Depends on topic, it would have its place</td>
</tr>
<tr>
<td>Would it be slow? (x2)</td>
<td>Good for broad topic (e.g. ethics x2)</td>
</tr>
<tr>
<td>Very keen to have it</td>
<td>Would it be slow? (neg react x2)</td>
</tr>
<tr>
<td>Quite often relevant information can be hidden away in surprisingly obscure places</td>
<td>Great/good 3</td>
</tr>
<tr>
<td>Sometimes different databases have different ways of interfacing – a standard search would be easier</td>
<td>Would prefer just medicine, metasearch too broad</td>
</tr>
</tbody>
</table>

The concept of federated searching was met with a positive response from both groups of participants. Several participants commented on the tedium and frustration of the status quo – the necessity to repeat searches in several databases and the difficulty of knowing which databases to use being the main reasons for frustration.

Interestingly, despite the reaction to the earlier questioning about the speed of current databases available to participants, the initial reaction of four participants was to ask whether such a search would be slow. Some went so far as to say that if the federated search was slower than current searches, they would not bother to use it.

The medical cohort, while being interested in the concept of a federated search on the whole, displayed more caution than the history cohort. While half of the medical cohort reacted with enthusiasm, some commented that while a federated search across all subjects would be useful, the level of usefulness might be limited. Two medical participants commented that a federated search would be more useful for a broad topic, for example ethics, than their more scientific topics. Another medical participant
commented that he felt that medicine was already a very broad topic, and providing an even broader search would not be useful to him.

On the other hand, the reaction of the history cohort was much more positive. Two history participants queried, unprompted, whether such a search would be slow. All other participants responded enthusiastically to the concept:

‘quite often relevant information can be hidden away in surprisingly obscure places’ – History participant #2

‘sometimes different databases have different ways of interfacing. A standard search would be easier’ – ibid.

‘very useful... instead of having to do four or five searches across the different areas... that would be very helpful’ – History participant #1

It is worth noting that this question related to a search that searched across all subjects, a ‘one box’ search. One or two of the earlier participants interviewed commented, unprompted, that a federated or faceted search that searched across their subject only would be more useful to them than a federated search that searched across all subjects. Following these early unprompted suggestions of a faceted search, all participants were asked about this option if they did not mention it spontaneously. All participants who were asked about a faceted option agreed that it would be useful to have.

Results from the final question (‘Is there anything else you would like to tell me about how you search for information?’) will be covered in the discussion section below. Some results were coded with the responses to the questions above, if they contributed further to the understanding of themes and concepts discussed earlier in the interview.
10 Discussion

What are the information needs of history and medical students when searching online databases?

The information needs of history and medical students according to this study are fairly straightforward. With such a small sample size it is not possible to generalise, but the results of this study give an indication of differences in attitudes of students of the two faculties towards federated searching, and provide scope for further study.

Students of both disciplines need to be able to find information that will answer their assignment question or topic, whatever that may be, reasonably quickly and with a minimum of fuss. As the results have shown, there is no delineated difference in the basic methods students of both disciplines use to find their resources. Several comments were received from participants about the tedium and frustration of the status quo, which requires students to search multiple databases and interfaces to find the information that they need. One participant went so far to say that she looks for articles that will ‘get me marks’.

Participants of both subjects usually have a ‘favourite’ database, one they regularly use or start with. Most students have a typical individual search process that they follow, developed from prior successful experiences. This may include external sites or resources such as Google, Google Scholar or other resources such as encyclopaedias. What did become clear from the responses received from participants was that there was no ‘typical’ search process followed by the majority of participants. While all participants used library databases for their research, the point at which a participant chose to use a database in the searching process varied, with some preferring to start their search for information elsewhere to confirm their understanding of the topic, or to refine their keywords, while others went directly to the databases to begin their searches.
Based on responses received from several participants, the typical search process referred to above relates to what participants are familiar and comfortable with. Familiarity and comfort are two concepts that reverberated throughout the interviews. While participants were not questioned directly about whether they were comfortable (or not) with their search process and results, many commented on this unprompted. Some commented that they were cautious about trying to use an unfamiliar interface (‘fumbling’ was a term used to describe this by one participant), mainly from a perspective of limited time available. This concurs with the comment made by Reeb (2006, p.353) about a user who preferred to continue with what he was familiar with, avoiding an ‘investment of time’ although other products may be better suited to his needs.

The issue of time investment is important, as it affects user behaviour. As Tennant (2007) stated in his Keynote speech and the Library and Information Association of New Zealand Aotearoa (LIANZA) conference, ‘Our users have lives’. It is all too easy for library and information professionals to forget this, and in the opinion of the researcher, library interfaces have been designed for the ‘back-end’ of the library (necessarily) for too long, with the user ‘front-end’ often working around various idiosyncrasies. The result has been a poor user experience, and the level of emotion elicited in response to some of the questions demonstrates this.

Searching can be an emotional process (Kuhlthau, 1991), and the number of comments received from participants relating to emotions bear this out. Several comments were made by participants about the tedium and frustration of the article/database searching process.

‘I think the searching process is quite a tedious process...’ – Medical participant #6

‘...it’s such a hassle having to go [to multiple databases], it’s not, I don’t find it particularly user-friendly’ – History participant #4

It is apparent from the comments above that these students do not enjoy the searching process, perhaps unsurprisingly. It is also clear that there is a gap in the provision of searching functionality that federated searching is attempting to fill.
Many participants commented that they did ‘what worked’ for them to find the resources they required, but felt that they were not searching the ‘best’ way. Some went so far as to say that they felt they needed more help with this aspect of their studies – slightly surprising, as this was not a line of enquiry that the interviewer followed up specifically.

‘I do think that my searching skills are not really efficient enough. I need to learn a lot more’ – Medical participant #3

Interestingly, more comments about how participants felt about their searching ability were received from medical participants than history participants. One medical participant mentioned that she had done a library-run course on the Ovid platform, recognising that she didn’t feel she was getting all she could have out of the interface. Two other medical participants made reference to search techniques that they were taught in class. History participants referred to particular databases recommended by lecturers or tutors, but only one history participant specifically mentioned being taught how to use a database during classes. No history participants mentioned attending library-run courses.

If participants were faced with an ‘unsuccessful’ search, i.e. irrelevant, or no results, they employed a variety of techniques to improve their searches. These included refinement of their search terms, judicious use of Google, use of the advanced functions in a particular database, or occasionally asking a librarian for help. The way in which these participants use Google demonstrates, for these students at least, that there is no need for librarians to have concerns about the way Google is being used for academic research. By using Google to confirm or refine the understanding of a topic, to check citations or find alternative keywords before going back to a library database, this cohort demonstrated a good understanding of the limitations and advantages of using Google.

How can the type of search result (e.g. precise ‘hit’ vs. broader result) affect the effectiveness of a search?
Responses to this line of questioning were fairly uniform across both subjects, in the sense that there was no discernable pattern. The interview question that most closely related to this research question was the question asking students whether they had a preference for specific or broad results. Responses to this question did not fall into a pattern delineated by subject, but the preference for a specific or broad result list may have more to do with individual learning styles than the subjects that participants study.

All participants seemed to be aware of the issue of relevance, and the potential of missing out on relevant information through ‘not knowing where to look’. Most participants were aware of and used advanced search techniques to refine their searches within a database.

While no correlation was found between the preferred type of results desired and the subject studied, a correlation was found between the year of study and the type of results. This result is not for the cohort as a whole – demographic information was not asked for directly. Most participants did volunteer their year of study, and of those that did, there was a noticeable trend for those who were further through their studies to prefer a more specific result set. Some also commented that they would have preferred a broader result in first year. The requirement for more specific results may well be related to the more specific nature of the assignments that are given out towards the end of an undergraduate degree, a natural progression from the more generalist assignments of first year.

These findings have implications for information literacy and user needs. The requirement for more specific results as students progress through their studies suggests undergraduate information literacy programmes may need to be specifically tailored for those (for example) in the third year of a three year programme, or perhaps a separate first year teaching programme. The information literacy programme for the medical cohort seems to be working well, several participants mentioned library teaching they had had, and one talked in detail about the training they were given for PubMed as part of their normal lectures.
What are the attitudes of history and medical students to a product that will search more than one database simultaneously, given the limitations of this type of searching? Is there a correlation between the subject studied and the type of search preferred?

The overall attitude from all participants to the concept of a ‘one-box’, all subjects federated search was generally positive, with just over half of those interviewed stating that such a federated search would be good to have. Participants are very aware of the limitations of the current searching methods available to them, and the need to become familiar with several different interfaces is a stumbling block for some:

‘I’d probably want to try multiple [searches] through one database, instead of fumbling with something I wasn’t used to’ – Medical participant #2

‘If I’m feeling brave’ – History participant #2 in response to questioning about his use of advanced search functionality.

The medical participants were less enthusiastic about the concept of the one box search on the whole. While half the cohort responded positively initially, the number of comments and further queries was in contrast to those of the history cohort. The comments received focussed on the speed of the search (consistent with the history cohort), and the way a federated search might affect the relevance of the results received.

The generally higher level of enthusiasm displayed by the history cohort to federated searching may relate to the fact that JSTOR, one of the more popular databases mentioned by the history cohort, is effectively a federated search. The JSTOR index covers a very wide range of databases and subjects already, so these users are already familiar with the idiosyncrasies of a much broader search index.

It is interesting to note the slightly paradoxical responses received from participants relating to the issue of database speeds. When questioned about whether the speed of a database would affect their decision to move on, most participants said it wasn’t an issue unless the database in question was very slow. Nonetheless, it was striking to note that the majority of
queries received from participants of both subjects about federated searching related to the possibility that such a search might be prohibitively slow.

The option to facet or restrict the databases searched with a federated search met with a more positive response from the medical cohort than the ‘one box’ concept. The history cohort were also enthusiastic about the concept of faceted searching, although one history participant brought up the issue of ‘censorship’ while discussing faceted searching and how it might work. He felt that a history faceted search might be too restrictive, as history as a subject can be fairly interdisciplinary by nature, and was worried that relevant information might be missed through using this type of search. This participant felt that a facet at the faculty level (e.g. Arts) might be more appropriate.

What did come out of this questioning was that there is unlikely to be a ‘one size fits all’ solution for a federated search at the institution, rather than the subject level. Some users prefer to be able to search across a very wide range of subjects, and others prefer to search at a much more local level. Enabling users to select their own level of specificity, as it were, will ensure the success of such a search interface. The ability to ‘jump’ out of a federated search into a particular database will allow users to access the advanced functionality that may not be available within the federated search. As the results of this study have shown, the medical cohort in particular, and all users in general make some use of advanced search functionality. While the advanced search function in a federated search may replicate the functionality of an ‘average’ advanced search, it is unlikely to cope with the level of specificity provided by the Medline search interface. A multi-layered approach to a federated search, perhaps allowing users to ‘drill down’ to more restrictive searches will cater for most of the concerns raised by this cohort.
11 Directions for further research

There are several options for further research. This study focused on only two subject areas, it would be possible to replicate this study with other subject areas and investigate if the results correlate.

Research that focuses on how students search as they progress through their studies would also be useful. The results of this study indicate that there may well be differences in the searching methods students use and the results they prefer as they progress through their studies. There is also potential for a wider study comparing the experiences of undergraduate and postgraduate students. Another interesting angle to investigate would be the attitude of academic staff to federated searching.

In terms of the practical aspect of setting up a federated search at an academic institution, usability testing with focus groups of students is essential. The lack of facilities and a live site to test on restricted this study to a broader, exploratory focus.

Research outside of the tertiary and public library sector would also be valuable. With the advent of hosting services, technology that has traditionally been restricted to larger libraries due to the resources required to run them may become viable for smaller, specialist libraries. A federated search could have application in the corporate sector, and the special library sectors for example.
12 Conclusion

The results of this study have shown that, while there are differences in the preferred searching methods of participants in this study, the differences in methods used between the two subject groups studied are negligible. There are also differences in the results preferred by the participants, but again, the difference in the type of results preferred (‘specific’ or ‘broad’) does not correlate to the subject studied. The differences in results and searching methods is more likely to be related to individual learning styles than the subject a student studies. Although demographic information was not collected from all participants, results from those who did indicate their year of study suggest that the level of specificity required from results of database searching increases as their studies progress. This correlates with other studies done in this area. These findings have implications for both information literacy programmes at academic tertiary institutions and implementation of a federated search product, although more detailed research will need to be undertaken on both of these topics.

The results of this study have also shown that, while participants were enthusiastic about the concept of a federated search across all databases, many had concerns. These concerns related to the speed of such a search, and the relevance of the results returned. Further questioning about a faceted or more restrictive searching option produced a much more positive response.

Situations where federated searching would be appropriate for the cohort of students who participated in this study include one where the option to facet a search, either by faculty, or at a departmental level was available. Some participants also expressed interest in being able to choose not to use a federated search, particularly the medical participants. One or two medical participants felt that a search, even faceted at the ‘Medicine’ level, would still be too broad, and preferred to use individual databases the same way they do now. Other medical participants commented that how useful a federated search was would depend on the particular assignment or topic they were studying at the time.
A ‘one box’ federated search is undoubtedly better than the status quo, and would be received with enthusiasm by some students. However, for a federated search to be appropriate for use in an academic tertiary setting, it is important to ensure user needs are met. As has been shown by this study, this cohort had a range of needs in relation to federated search. The most important of these are the option to facet a search, and a searching speed that is not prohibitively slow.
13 Bibliography


14 Appendices

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.1</td>
<td>Participant Information Sheet</td>
<td>37</td>
</tr>
<tr>
<td>14.2</td>
<td>Participant Consent form</td>
<td>38</td>
</tr>
<tr>
<td>14.3</td>
<td>Interview questions:</td>
<td>39</td>
</tr>
<tr>
<td>14.4</td>
<td>Sample transcript</td>
<td>40</td>
</tr>
<tr>
<td>14.5</td>
<td>Sample notice from Cecil</td>
<td>41</td>
</tr>
<tr>
<td>14.6</td>
<td>Sample poster</td>
<td>42</td>
</tr>
</tbody>
</table>
14.1 Participant Information Sheet

Dear Participant,

My name is Vanessa Newton-Wade, and I am studying towards my Masters in Library and Information Studies at Victoria University of Wellington.

For my final assessment, I am required to carry out research on an aspect of the library sector. The objectives of the research I am conducting are to identify situations where the ability to search across a wide range of databases from all subjects simultaneously may be appropriate for academic tertiary libraries, and to find out whether students from different academic disciplines prefer to search in the same way or not.

Participating in the research will take no more than 15-20 minutes, and will involve an interview. The interview will involve 6 set questions, some searching on a library database you are familiar with, and time for further discussion.

Interviews will be recorded using a digital voice recorder and transcribed. You will not be identified by name in any part of the final report, and access to the recordings and transcripts will be restricted to my supervisor and a research assistant. The data will be kept confidential, and destroyed one year after the completion of the project.

Should you wish to pull out of the research after completing the interview, you may do so within 2 weeks of the interview date, without giving a reason. If you are interested in the results of the research, I am happy to provide you with a summary of these.

This research will be written up in the form of a report. A copy of the final report will be deposited in the VUW library, and also online in the VUW institutional repository. It may also be submitted for publication in academic journals.

If you have any questions or concerns, please feel free to contact myself, or my supervisor. Details are at the end of this sheet.

Thank you for your time and interest in participating in this project! It would not be possible without you.

Vanessa Newton-Wade
Masters Student (Masters in Library and Information Studies)
School of Information Management
Victoria University of Wellington

Email newtonvane@vuw.ac.nz
Phone (09) 373 7599 x82154

Brenda Chawner (Supervisor)
School of Information Management
brenda.chawner@vuw.ac.nz
04 463 5780
14.2 Participant Consent form

I, _______________________________________________________________, have read and understood the enclosed participant information sheet, and give my consent to be involved in this project.

I understand that

- I will not be identified by name in the final report
- My interview recording and transcript will be confidential to the research team
- My interview recording and transcript will be destroyed after 1 year
- I may request a summary of the findings of this research
- I may request to view a copy of my interview transcript within 2 weeks of the interview date.
- I may withdraw from the project, without giving a reason, at any time up until 2 weeks after the interview date.

☐ Yes, please contact me with a summary of the research findings

☐ Yes, I wish to see a copy of my interview transcript

Signed _______________________________________________________________

Date ______________

Please write your contact details below if you wish to receive a summary of the research findings and/or a copy of your interview transcript.
14.3 Interview questions:
Starting point: Think about, or imagine, beginning research on a topic you know little about, using a library database you are familiar with. You are not searching for known references (e.g. an article or author you have a reference for).

An example might be beginning to look for information on an essay topic.

1. How do you go about your initial search?
2. What do you consider a successful search?
3. What would you do differently if your search is unsuccessful?
4. What is your tolerance limit before 'giving up' on a database and trying something else? What are your typical reasons for ‘giving up’ on a database? (Potential probing questions will relate to – speed of database, type / relevance / lack of results, etc.)
5. Do you have a preference for results that are specific or broad in terms of the number of results returned and their relevance, and why? How specific do you need the information to be at this point in your research? Bear in mind that you are just beginning the research.
6. What is your reaction to a search engine that can search all of the library databases from all subjects at once? (potential probing questions relate to – relevance / speed of search etc.)
7. Is there anything else you would like to tell me about how you search for information?

The observation component of the interviews is intended to allow students to conduct a search in a database they are familiar with, to assist with answering the questions. This allows students some context, rather than answering the questions ‘cold’. The database and search terms used will be noted – results will be discussed in relation to the results-based questions above, and probing questions may be asked as students search.
14.4 Sample transcript

V: so when you're in the database searching, what would you consider a successful search? Just thinking that this is the beginning point, you've gone out and looked at your books and you're just starting to look at the electronic stuff.

I: A successful search would probably be one or two, at least one full-text piece, that deals quite directly with the subject topic, that offers quite a lot specifically about the essay. Rather than just some random...

V: and what would you do differently if you didn’t find what you wanted?

I: I’d rework my search terms, see if they had anything to do with it, and then often, if I found one good article, then I’ll go through and look at the bibliography for that. And then start to look for those kinds of things.

V: And what would your tolerance limit be before you gave up on a database and tried something else, whether it was another database, or

I: probably the longest I’ve ever been on a database would be for an hour, which seems horrible and really exhausting, I dread, I actually don’t look forward to having to go through the databases and article searching process, because that just feels overwhelming sometimes. And you forget, say for history for example, when we do book reviews, so you need to find sources for two different things, and then you lose track of every site you’re going through because you’re doing the title twice, yeah, it gets really confusing. I think the most on one topic would probably be an hour, and then I need to have a break and ... go away. On one site, anyway.

V: would the speed of a database have an effect on your tolerance limits?

I: Um, Definitely [emphasis], the library a little while ago was having a problem with JSTOR, they couldn’t access it. That completely throws you off, and they had people asking for extensions and things because it affects your work so greatly. If you’re sitting there for twenty minutes just waiting for it to bring up one result, it’s just terrible. Yes, speed is a big factor.

V: when you’re looking for information for your essays and things, do you have a preference for results that are specific, or broad?

I: How do you mean?

V: Basically what I’m looking for is, um, when you’re at this beginning point in looking for information, are you looking for something really really targeted, or do you prefer to read around the edges?

I: I see. I prefer something targeted, because I’m not normally in a position where I have lots of time to spend doing broad reading, I think you develop the broad reading... I guess I could go for specific, and then if you have lots of different specific elements, like things that focus on different specific things within one topic, then you end up with a broader knowledge, but I wouldn’t... But having said that, broad pieces that have specific parts to them can be really helpful, like chronologies. But I’d prefer specific over broad. With history especially you don’t tend to get questions which are so substantive that they’re, that you need a broad thing.

V: and what would your reaction be to a search engine that could search all of the library databases across all subjects at once?

I: that’d be great! Yeah, that’d be fantastic. Because it’s such a hassle having to go, especially with these (gestures), it’s not, I don’t find it particularly user-friendly. If you go to databases, unless you know the specific title that you’re looking for, otherwise you have to search by subject and then even then it doesn’t really tell you... it’s not interdisciplinary. I’ll be looking for... I’m doing a European Film paper at the moment, and instead of going into European lang and lit, I’d go to FMTV because this one doesn’t have that much to offer in terms of European tv so ... yeah.
Got 10 minutes to spare? Want a free coffee/hot chocolate?

Med students needed for research

Hi everyone!

I am still looking for volunteers from years 2-5 for short interviews about library database searching, for my final Masters project.

If anyone can help me out I’d really appreciate it!! Interviews will take about 10-15 minutes.

I will be in the Seminar Room at the Philson Library this Thursday (29th May) from 1-2pm, please drop in!

Or contact me by email (newtonvane@student.vuw.ac.nz) to arrange a time/place that suits you.

If you have any queries about the research, please contact me, or my supervisor (details at end).

Thanks all!!

Cheers,
Vanessa

Vanessa Newton-Wade
Masters Student, Library and Information Studies
Victoria University of Wellington

Email newtonvane@student.vuw.ac.nz
Phone 09 373 7599 x82154

Supervisor: Brenda Chawner
Email brenda.chawner@vuw.ac.nz
Phone 04 463 5780
History students needed!!

Are you a 2nd/3rd yr History Student?
Do you use library databases for research?

I am conducting research for my Masters in Library Studies and am looking for history students to interview about database use.

Please contact Vanessa if you are interested in participating!

The Fine Print: Interviews take about 30 min, are conducted on campus & chocolate fish are offered as a token of thanks ☺